```
OM nucleic - nucleic search, using sw model

                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:             September 20, 2004, 15:12:19 ; Search time 147.339 Seconds
                                                   (without alignments)
                                                   7871.356 Million cell updates/sec

Title:              US-09-825-769A-12
Perfect score:      273
Sequence:           1 cctggtgtcggtgatgatgg..........cacggcggcggccacgaacg 273
Scoring table:      IDENTITY NUC
                    Gapop 10.0 , Gapext 1.0
Searched:           3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters:        6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing:    Minimum Match  0%
                    Maximum Match 100%
                    Listing first 45 summaries

Database :          N_Geneseq_29Jan04:*
                    1:  geneseqn19808:*
                    2:  geneseqn19908:*
                    3:  geneseqn20008:*
                    4:  geneseqn2001as:*
                    5:  geneseqn2001bs:*
                    6:  geneseqn20028:*
                    7:  geneseqn2003as:*
                    8:  geneseqn2003bs:*
                    9:  geneseqn2003cs:*
                   10:  geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

Result          %     Query
No.    Score  Match  Length  DB  ID          Description
-----  -----  -----  ------  --  ----------  -----------
   1    272    99.6    273    6  Aas15486    Aas15486 B. pertus
   2    270.4   99.0  1212    6  Aas15484    Aas15484 DNA encod
   3    189.4   69.4  1212    6  Abq90070    Abq90070 M. capsul
   4    153.6   56.3  1215    6  Aaz09786    Aaz09786 E. coli b
   5    153.6   56.3  1232    2  Aax02811    Aax02811 E. coli b
   6    153.6   56.3  1419    5  Aas88509    Aas88509 DNA encod
 c 7    153.6   56.3  1548    5  Aas88547    Aas88547 DNA encod
   8    153.6   56.3  1548    5  Aas93345    Aas93345 DNA encod
   9    153.6   56.3  1548    5  Aas73075    Aas73075 DNA encod
 c 10   153.6   56.3  3414    2  Aas73033    Aas73033 DNA encod
 c 11   153.6   56.3  3481    2  Aax02815    Aax02815 DE1973127
 c 12   153.6   56.3  3810    2  Aax02813    Aax02813 DE1973127
   13   153.6   56.3  10614   4  Aas46259    Aas46259 DNA encod
 c 14   148.8   54.5  1212    7  Abz39842    Abz39842 N. gonorr
 c 15   140.8   51.6  13786   3  Aaa81480    Aaa81480 N. mening
 c 16   140.8   51.6  110000  3  Aaa81490_13 Continuation (14 o
   17   140.8   51.6  349980  3  Aaf21610    Aaf21610 Neisseria
   18   139.4   51.1  425     7  Aas15487    Aas15487 B. pertus
   19   120.6   44.2  1215    7  Acf72052    Acf72052 Photorhab
   20   120.6   44.2  110000  7  Acf67367_52 Continuation (53 o
   21   120.6   44.2  110000  7  Acf65387_4  Continuation (5 of
   22   117.4   43.0  1518    4  Abl09101    Abl09101 Drosophil
   23   117.4   43.0  3041    4  Abl11656    Abl11656 Drosophil
   24   117.4   43.0  3576    4  Abl09100    Abl09100 Drosophil
   25   112     41.0  62909   4  Aaf28545    Aaf28545 Genomic f
   26   111     40.7  1140    7  Abt18306    Abt18306 Aspergill
   27   111     40.7  1140    7  Abt18900    Abt18900 Aspergill
   28   111     40.7  1554    7  Abt20720    Abt20720 Aspergill
   29   111     40.7  1778    7  Abt20122    Abt20122 Aspergill
   30   111     40.7  3140    7  Abt17712    Abt17712 Aspergill
   31   111     40.7  3778    7  Abt19526    Abt19526 Aspergill
   32   106.2   38.9  1260    8  Ada32431    Ada32431 DNA encod
   33   105.8   38.8  565     1  Aas15485    Aas15485 B. pertus
 c 34   90.8    33.3  110000  2  Aat42063_03 Continuation (4 of
   35   88.8    32.5  1362    6  Abz12797    Abz12797 Arabidops
   36   88.8    32.5  1399    7  Aaz98363    Aaz98363 A. thalia
   37   88.8    32.5  1494    7  Aad48348    Aad48348 Saccharom
   38   88.8    32.5  1656    3  Aac42345    Aac42345 Arabidops
   39   88.8    32.5  11198   4  Aad07496    Aad07496 PYLR110P
 c 40   88.8    32.5  11201   4  Aad07498    Aad07498 PYMR107P
 c 41   88.8    32.5  11204   4  Aad07499    Aad07499 PZEO1P DN
 c 42   88.8    32.5  11427   4  Aad07497    Aad07497 PYMR251AP
 c 43   88.8    32.5  12008   4  Aad07500    Aad07500 PPR1BP DN
 c 44   88.8    32.5  12644   4  Aad07492    Aad07492 PYLR110P+
   45   88.8    32.5  12850   4  Aad07495    Aad07495 PZEO1P+1u
```

                              ALIGNMENTS

```
RESULT 1
AAS15486
ID   AAS15486 standard; DNA; 273 BP.
XX
AC   AAS15486;
XX
DT   29-AUG-2003  (revised)
DT   14-FEB-2002  (first entry)
XX
DE   B. pertussis cysteine desulphinase DNA (DSF536R1) from strain BP536.
XX
KW   whooping cough vaccine; cysteine desulphinase; ds.
KW   Large-scale bacterial toxin production; pertussis toxin; PT;
OS   Bordetella pertussis; strain BP536.
XX
PN   WO200174862-A2.
XX
PD   11-OCT-2001.
XX
PF   04-APR-2001; 2001WO-US010938.
XX
PR   04-APR-2000; 2000US-0194478P.
PR   04-APR-2000; 2000US-0194482P.
XX
PA   (BAXT ) BAXTER INT INC.
PA   (BAXT-) BAXTER HEALTHCARE SA.
XX
PI   Blake MS, Bogdan JA, Nazario-Larrieu J;
XX
DR   WPI; 2002-010777/01.
XX
PT   Enhancing production of bacterial toxins comprises eliminating or
PT   reducing toxin expression inhibitors formed by toxin producing bacteria
PT   by adding at least one soluble metal salt that forms an insoluble complex
PT   with sulfate.
XX
PS   Disclosure; Fig 7; 46pp; English.
XX
CC   The present invention relates to a method of enhancing the production of
CC   bacterial toxins in large-scale culture. The method comprises
CC   eliminating or reducing toxin expression inhibitors formed by toxin
CC   producing bacteria. The invention provides a method for producing
CC   pertussis toxin (PT) from Bordetella pertussis by introducing a soluble
CC   salt into the growth medium that sequesters sulphate and/or employing a
CC   B. pertussis cysteine desulphinase knockout mutant. The method is useful
```

```
OM nucleic - nucleic search, using sw model

                        GenCore version 5.1.6
                        Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:                 September 20, 2004, 15:12:19 ; Search time 229.374 Seconds
                                                       (without alignments)
                                                       7871.356 Million cell updates/sec

Title:          US-09-825-769A-11
Perfect score:  425
Sequence:       1 ctcgacctgcagaagctgaa................ggttcggcctgcacgtcggc  425
Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched:       3373863 seqs, 212409041 residues
Total number of hits satisfying chosen parameters:   6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      N_Geneseq_29Jan04:*
           1:   geneseqn1980s:*
           2:   geneseqn1990s:*
           3:   geneseqn2000s:*
           4:   geneseqn2001as:*
           5:   geneseqn2001bs:*
           6:   geneseqn2002s:*
           7:   geneseqn2003as:*
           8:   geneseqn2003bs:*
           9:   geneseqn2003cs:*
          10:   geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

```
Result            %        Query
No.     Score   Match  Length  DB  ID        Description
------  -----   -----  ------  --  --        -----------
     1  425     100.0   425     6  AAS15487  Aas15487 B. pertus
     2  424      99.8  1212     6  AAS15484  Aas15484 DNA encod
     3  334      78.6   560     6  AAS15488  Aas15488 B. pertus
     4  288      67.8  1212     6  ABQ90070  Abq90070 M. capsul
     5  246.4    58.0  1215     2  AAZ09786  Aaz09786 E. coli b
     6  246.4    58.0  1232     2  AAX02811  Aax02811 E. coli b
     7  246.4    58.0  1548     5  AAS88547  Aas88547 DNA encod
     8  246.4    58.0  1548     5  AAS93345  Aas93345 DNA encod
     9  246.4    58.0  1548     5  AAS73075  Aas73075 DNA encod
  c 10  246.4    58.0  3414     5  AAS73033  Aas73033 DNA encod
    11  246.4    58.0  3481     2  AAX02815  Aax02815 DE1973127
    12  246.4    58.0  3810     2  AAX02813  Aax02813 DE1973127
    13  246.4    58.0 10614     4  AAS46259  Aas46259 DNA encod
    14  232.6    54.7  1212     7  ABZ39842  Abz39842 N. gonorr
  c 15  213.4    50.2 13786     3  AAA81480  Aaa81480 N. mening
  c 16  213.4    50.2 110000     3  AAA81490_13 Aaa81490_13
  c 17  213.4    50.2 349980     3  AAF21610  Aaf21610 Neisseria
    18  184      43.3  1215     7  ACF72052  Acf72052 Photorhab
    19  184      43.3 110000     7  ACF67367_52 Acf67367_52
    20  184      43.3 110000     7  ACF65387_4  Continuation (5 of
  c 21  162      38.1  1419     5  AAS88509  Aas88509 DNA encod
    22  158.4    37.3 62909     4  AAF28545  Aaf28545 Genomic f
    23  156.8    36.9  1260     8  ADA32431  Ada32431 DNA encod
```

```
  c 24  147.2    34.6  1554     7  ABT20720  Abt20720 Aspergill
    25  144.8    34.1 110000     2  AAT42063_03 Continuation (4 of
    26  139.4    32.8   273     6  AAS15486  Aas15486 B. pertus
    27  119.2    28.0  1518     4  ABL09101  Abl09101 Drosophil
    28  119.2    28.0  3041     4  ABL11656  Abl11656 Drosophil
    29  119.2    28.0  3576     4  ABL09100  Abl09100 Drosophil
    30  118.4    27.9  1434     9  ADB69794  Adb69794 C. neofor
    31  118.4    27.9 40680     6  ABA92787_6  Continuation (7 of
    32  112.8    26.5  1649     3  AAF18102  Aaf18102 Lung canc
    33  112.8    26.5  2380     4  AAH14339  Aah14339 Human CDN
    34  112.8    26.5  3224     4  ABV27901  Abv27901 Human pro
    35  109      25.6  1362     6  ABZ12797  Abz12797 Arabidops
    36  109      25.6  1399     3  AAZ98363  Aaz98363 A. thalia
    37  109      25.6  1656     9  AAC42345  Aac42345 Arabidops
    38  107      25.2  1653     9  ADB69433  Adb69433 C. neofor
    39  107      25.2  3653     9  ADB69072  Adb69072 C. neofor
    40  102.4    24.1  2092     4  AAH14100  Aah14100 Human CDN
    41  97.2     22.9  1140     7  ABT18306  Abt18306 Aspergill
    42  97.2     22.9  1140     7  ABT18900  Abt18900 Aspergill
    43  97.2     22.9  1778     7  ABT20122  Abt20122 Aspergill
    44  97.2     22.9  3140     7  ABT17712  Abt17712 Aspergill
    45  97.2     22.9  3778     7  ABT19526  Abt19526 Aspergill
```

ALIGNMENTS

```
RESULT 1
AAS15487
ID    AAS15487 standard; DNA; 425 BP.
XX
AC    AAS15487;
XX
DT    29-AUG-2003  (revised)
DT    14-FEB-2002  (first entry)
XX
DE    B. pertussis cysteine desulphinase DNA (DSF53611) from strain BP536.
XX
KW    Large-scale bacterial toxin production; pertussis toxin; PT;
KW    whooping cough vaccine; cysteine desulphinase; ds.
XX
OS    Bordetella pertussis; strain BP536.
XX
PN    WO200174862-A2.
XX
PD    11-OCT-2001.
XX
PF    04-APR-2001; 2001WO-US010938.
XX
PR    04-APR-2000; 2000US-0194478P.
PR    04-APR-2000; 2000US-0194482P.
XX
PA    (BAXT ) BAXTER INT INC.
PA    (BAXT-) BAXTER HEALTHCARE SA.
XX
PI    Blake MS, Bogdan JA, Nazario-Larrieu J;
XX
DR    WPI; 2002-010777/01.
XX
PS    Disclosure; Fig 7; 46pp; English.
XX
CC    The present invention relates to a method of enhancing the production of
CC    bacterial toxins in large-scale cultures. The method comprises
CC    eliminating or reducing toxin expression inhibitors formed by toxin
CC    producing bacteria. The invention provides a method for producing
CC    pertussis toxin (PT) from Bordetella pertussis by introducing a soluble
CC    salt into the growth medium that sequesters sulphate and/or employing a
CC    B. pertussis cysteine desulphinase knockout mutant. The method is useful
```

Enhancing production of bacterial toxins comprises eliminating or
reducing toxin expression inhibitors formed by toxin producing bacteria
by adding at least one soluble metal salt that forms an insoluble complex
with sulfate.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2004, 15:12:20 ; Search time 1213.07 Seconds
                 (without alignments)
                 9754.272 Million cell updates/sec

Title: US-09-825-769A-12
Perfect score: 273
Sequence: 1 cctggtgtcggtgatgatgg.........cacggcggcggccacgaacg 273

Scoring table: IDENTITY NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167516995 residues

Total number of hits satisfying chosen parameters: 6940544

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :   GenEmbl:*
                 1: gb_ba:*
                 2: gb_htg:*
                 3: gb_in:*
                 4: gb_om:*
                 5: gb_ov:*
                 6: gb_pat:*
                 7: gb_ph:*
                 8: gb_pl:*
                 9: gb_pr:*
                 10: gb_ro:*
                 11: gb_sts:*
                 12: gb_sy:*
                 13: gb_un:*
                 14: gb_vi:*
                 15: gb_ba:*
                 16: em_fun:*
                 17: em_hum:*
                 18: em_in:*
                 19: em_mu:*
                 20: em_om:*
                 21: em_or:*
                 22: em_ov:*
                 23: em_pat:*
                 24: em_ph:*
                 25: em_pl:*
                 26: em_ro:*
                 27: em_sts:*
                 28: em_un:*
                 29: em_vi:*
                 30: em_htg_hum:*
                 31: em_htg_inv:*
                 32: em_htg_other:*
                 33: em_htg_mus:*
                 34: em_htg_pln:*
                 35: em_htg_rod:*
                 36: em_htg_mam:*
                 37: em_htg_vrt:*
                 38: em_sy:*
                 39: em_htgo_hum:*
                 40: em_htgo_mus:*
                 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		270.4	99.0	346274	1	BX640443	BX640443 Bordetell
2		270.4	99.0	344954	1	BX640416	BX640416 Bordetell
3		268.8	98.5	344321	1	BX640429	BX640429 Bordetell
4		220.8	80.9	193050	1	AL646062	AL646062 Ralstonia
5		208	76.2	301846	1	AB016913	AB016913 Chromobac
6	c	188.8	69.2	14652	1	AB004799	AB004799 Pseudomon
7	c	184	67.4	7886	1	AP010139	AP010139 Azotobact
8		168	61.5	300861	1	AE016777	AE016777 Pseudomon
9	c	162.2	59.4	311249	1	AE016860	AE016860 Pseudomon
10		158.4	58.0	1251	6	AR387886	AR387886 Sequence
11		157.4	57.7	10107	1	AE004633	AE004633 Pseudomon
12	c	157.4	57.7	145050	1	AL627275	AL627275 Salmonell
13		157.4	57.7	300431	1	AE016835	AE016835 Salmonell
14	c	155.8	57.1	19982	1	AB005483	AB005483 Escherich
15	c	155.2	56.8	9356	1	AB008815	AB008815 Escherich
16	c	155.2	56.8	270365	1	AP002562	AP002562 Escherich
17		155.8	56.8	329861	1	NMA522491	NMA522491 Neisseria
18		153.6	56.3	1215	6	AX018964	AX018964 Sequence f
19		153.6	56.3	1215	6	BD136659	BD136659 Process f
20		153.6	56.3	1232	6	AX004472	AX004472 Sequence
21		153.6	56.3	1233	6	BD073376	BD073376 Process f
22		153.6	56.3	3481	6	AX000480	AX000480 Sequence f
23		153.6	56.3	3481	6	BD073380	BD073380 Process f
24		153.6	56.3	3810	6	AX000476	AX000476 Sequence
25		153.6	56.3	3810	6	BD073378	BD073378 Process f
26		153.6	56.3	9329	1	AE015273	AE015273 Shigella
27		153.6	56.3	10614	6	AB000339	AB000339 Escherich
28		153.6	56.3	10614	6	AX370241	AX370241 Sequence
29		153.6	56.3	13332	1	D90883	D90883 E.coli geno
30		153.6	56.3	241679	2	AC020861	AC020861 Mus muscu
31		153.6	56.3	290380	1	AE016987	AE016987 Shigella
32		153.6	56.3	300099	1	AE016764	AE016764 Escherich
33		143	52.4	1450	8	CRU279023	AJ279023 Candida r
34	c	141.4	51.8	308015	1	AE016783	AE016783 Pseudomon
35		140.8	51.6	10203	1	AE002486	AE002486 Neisseria
36		140.8	51.6	349980	6	AX044032	AX044032 Sequence
37		137	50.2	159119	1	AP006583	AP006583 Gloeobact
38	c	135.2	49.5	313518	1	AE016856	AE016856 Pseudomon
39	c	135	49.5	251720	1	AP005333	AP005333 Vibrio vu
40	c	135	49.5	301442	1	AE016798	AE016798 Vibrio vu
41	c	133.4	48.9	301235	1	AP005075	AP005075 Vibrio pa
42	c	130.2	47.7	310029	1	AE016868	AE016868 Pseudomon
43		128.6	47.1	6149	1	AY138456	AY138456 Xenorhabd
44		128	46.9	1605	8	AK059171	AK059171 Oryza sat
45		128	46.9	1816	8	AK072043	AK072043 Oryza sat

ALIGNMENTS

RESULT 1
BX640443
LOCUS   BX640443   346274 bp   DNA   linear   BCT 14-AUG-2003
DEFINITION Bordetella bronchiseptica strain RB50, complete genome; segment
            7/16.
ACCESSION   BX640443 BX470250
VERSION   BX640443.1 GI:33575370
KEYWORDS   complete genome.
SOURCE   Bordetella bronchiseptica RB50
  ORGANISM Bordetella bronchiseptica RB50
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Alcaligenaceae; Bordetella.
REFERENCE   1
  AUTHORS   Parkhill,J., Sebaihia,M., Preston,A., Murphy,L.D., Thomson,N.,
            Harris,D.B., Holden,M.T.G., Churcher,C.R., Bentley,S.D.,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2004, 15:12:20 ; Search time 993.414 Seconds
(without alignments)
8206.421 Million cell updates/sec

Title: US-09-825-769A-12
Perfect score: 273
Sequence: 1 cctggtgtcggtgatgatgg.........cacggcggcggccacgaacg 273

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 1493109276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
EST:*
    1:  em_estba:*
    2:  em_esthum:*
    3:  em_estin:*
    4:  em_estmi:*
    5:  em_estov:*
    6:  em_estpl:*
    7:  em_estro:*
    8:  em_htc:*
    9:  gb_est1:*
   10:  gb_est2:*
   11:  gb_est3:*
   12:  gb_est4:*
   13:  gb_est5:*
   14:  gb_htc:*
   15:  em_estfun:*
   16:  em_estom:*
   17:  em_esthum:*
   18:  em_gss_inv:*
   19:  em_gss_pln:*
   20:  em_gss_vrt:*
   21:  em_gss_fun:*
   22:  em_gss_mam:*
   23:  em_gss_mus:*
   24:  em_gss_pro:*
   25:  em_gss_rod:*
   26:  em_gss_phg:*
   27:  em_gss_vrl:*
   28:  em_gss_hum:*
   29:  gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	155	56.8	904	28	BZ554829	pacs1-60
	2	154.2	56.5	964	28	BZ573485	BZ573485 msh2_3220
c	3	148.6	54.4	1172	28	BZ558565	BZ558565 pa98401_2
c	4	147.8	54.1	602	9	AL932800	AL932800
	5	147	53.8	1202	28	BZ568162	BZ568162 pacs2-164
	6	143.8	52.7	715	13	BX613550	BX613550
	7	138.2	52.1	467	12	BI718906	BI718906 1031032H0
	8	138.2	50.6	607	12	BJ247332	BJ247332
	9	138.2	50.6	658	13	BU672399	BU672399 WHB3304_D
	10	137.6	50.4	378	11	W06757	W06757 SMRST0393_S
	11	135	49.5	980	11	CNS08MEH	CNS08MEH
c	12	132.4	48.5	853	28	BZ576834	BZ576834 Single re
	13	131.8	48.3	535	9	AV909852	AV909852 AV909852
	14	131.8	48.3	583	9	AV909567	AV909567 AV909567
	15	131.8	48.3	622	10	BE195251	BE195251 HVSMBH008
	16	131.8	47.5	640	11	CA014990	CA014990 HT12P02r
	17	129.6	47.4	429	10	BF483309	BF483309 WHB1791_C
	18	129.4	47.0	673	12	BI997818	BI997818 1031052A1
c	19	128.2	47.0	1148	28	BZ575496	BZ575496 msh2_4487
	20	128	46.9	407	28	AZ125543	AZ125543 OSjNBb007
	21	127	46.5	554	14	CF847417	CF847417 peHB043xF
	22	122.2	44.8	1491	11	AY107707	AY107707 Zea mays
	23	120.6	44.2	690	13	BX622505	BX622505
	24	120	44.0	571	10	BE593476	BE593476
	25	119.6	43.8	668	11	BG842680	BG842680 1024021B0
	26	119.5	43.4	941	11	BX018452	BX018452 Single re
	27	118.4	43.4	352	14	CD029570	CD029570 mgmk006xD
	28	118.2	43.3	781	14	CF672902	CF672902 RTCNT1_74
c	29	118.2	43.2	863	28	BZ562611	BZ562611 pacs2-164
	30	117.8	43.0	467	10	BF499916	BF499916 ATI4695.5
	31	117.4	43.0	729	12	BI354557	BI354557 GM27255.5
c	32	117	42.9	775	28	BZ560160	BZ560160 pacs2-164
	33	111.8	41.0	773	13	BX870153	BX870153
	34	111.2	39.3	843	14	CA789097	CA789097 AGENCOURT
	35	107.2	39.1	660	13	BW228002	BW228002
	36	106.8	39.1	674	13	CA071128	CA071128 SCCCAM100
	37	106.8	38.7	847	14	CF066829	CF066829 AcS17 Amp
	38	105.6	38.5	655	14	CA186265	CA186265 SCSGST311
	39	105.2	38.4	370	12	BI996051	BI996051 1031032H0
	40	104.8	37.9	639	13	BW226648	BW226648
	41	104.8	37.8	673	13	BW223613	BW223613
c	42	103.6	36.3	1156	28	BZ568061	BZ568061 pacs2-164
	43	102	36.1	692	13	CA071140	CA071140 SCCCAM100
	44	99	35.8	869	14	CB659834	CB659834 OSjNBd01B
	45	98.8	35.7	521	29	CC944467	CC944467 BOIPC94TF

ALIGNMENTS

```
RESULT 1
BZ554829
BZ554829 pacs1-60_4879-Y2 pacs1-60 Pseudomonas aeruginosa genomic clone

LOCUS       BZ554829          904 bp    DNA    linear   GSS 17-DEC-2002
DEFINITION  pacs1-60_4879-Y2 pacs1-60 Pseudomonas aeruginosa genomic clone
            pacs1-60_4879, genomic survey sequence.
ACCESSION   BZ554829
VERSION     BZ554829.1  GI:27162875
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
ORGANISM    Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 904)
AUTHORS     Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.B., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE       Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
JOURNAL     J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 206221695 4
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
```

OM nucleic - nucleic search, using sw model

Run on: September 20, 2004, 15:12:20 ; Search time 1888.48 Seconds
                (without alignments)
                9754.272 Million cell updates/sec

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Title: US-09-825-769A-11
Perfect Score: 425
Sequence: 1 ctcgacctgcagaagctgaa.........ggttcggcctgcacgtcggc 425

Scoring table: IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
                1: gb_ba:*
                2: gb_htg:*
                3: gb_in:*
                4: gb_om:*
                5: gb_ov:*
                6: gb_pat:*
                7: gb_ph:*
                8: gb_pl:*
                9: gb_pr:*
                10: gb_ro:*
                11: gb_sts:*
                12: gb_sy:*
                13: gb_un:*
                14: gb_vi:*
                15: gb_ba:*
                16: em_fun:*
                17: em_hum:*
                18: em_in:*
                19: em_mu:*
                20: em_om:*
                21: em_or:*
                22: em_ov:*
                23: em_pat:*
                24: em_ph:*
                25: em_pl:*
                26: em_ro:*
                27: em_sts:*
                28: em_un:*
                29: em_vi:*
                30: em_vrt:*
                31: em_hum:*
                32: em_htg_inv:*
                33: em_htg_other:*
                34: em_htg_mus:*
                35: em_htg_pln:*
                36: em_htg_rod:*
                37: em_htg_mam:*
                38: em_htg_vrt:*
                39: em_sy:*
                40: em_htgo_hum:*
                41: em_htgo_mus:*
                41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		424	99.8	349354	1	BX640416	Bordetell
2		422.4	99.4	346274	1	BX640443	Bordetell
3		417.6	98.3	344321	1	BX640429	Bordetell
4		329	77.4	193050	1	AL646062	Ralstonia
5	c	302.4	71.2	301846	1	AB016913	Chromobac
6	c	280	65.9	14652	1	AB004799	Pseudomon
7	c	260.8	61.4	7886	1	AF010139	Azotobact
8		254.4	59.9	1251	6	AR387886	Sequence
9	c	248	58.4	9356	1	AE005483	Escherich
10	c	248	58.4	270365	1	AP002562	Escherich
11		248	58.4	311249	1	AX018964	Sequence
12		246.4	58.0	1215	6	AX018964	Sequence
13		246.4	58.0	1215	6	BD136659	Process f
14		246.4	58.0	1232	6	AX000472	Sequence
15		246.4	58.0	1233	6	BD073376	Process f
16		246.4	58.0	3481	6	AX000480	Sequence
17		246.4	58.0	3481	6	BD073380	Process f
18		246.4	58.0	3810	6	AX000476	Sequence
19		246.4	58.0	3810	6	BD073378	Process f
20		246.4	58.0	10614	6	AX000339	Escherich
21		246.4	58.0	10614	6	BD073378	Process f
22		246.4	58.0	13332	6	AX370241	Sequence
23		246.4	58.0	241679	2	D90883	B.coli geno
24	c	246.4	58.0	9329	1	AC020861	Mus muscu
25	c	244.8	57.6	290380	1	AE015273	Shigella
26	c	244.8	57.6	299980	1	AE016987	Shigella
27	c	243.2	57.2	300099	1	AE016764	Escherich
28		240	56.5	150050	1	AL627275	Salmonell
29	c	240	56.5	304431	1	AE016835	Salmonell
30	c	238.4	56.1	19982	1	AE008815	Salmonell
31	c	235.2	55.3	300861	1	AE016777	Pseudomon
32	c	227.8	53.6	329861	1	NMA524991	Neisseria
33	c	213.4	50.2	102203	1	AX002486	Neisseria
34		213.4	50.2	349980	6	AX044032	Sequence
35	c	198.4	46.7	13584	1	AE013736	Yersinia
36	c	198.4	46.7	204050	1	AJ414154	Yersinia
37		193.6	45.6	251720	1	AP005333	Vibrio vu
38	c	190.4	44.8	301235	1	AP005075	Vibrio pa
39		188.8	44.4	14261	1	AE004160	Salmonell
40	c	188.8	44.4	301442	1	AE016798	Vibrio ch
41		186.2	43.8	11059	1	AE015668	Vibrio vu
42	c	185.6	43.7	10107	1	AE004633	Shewanell
43	c	184	43.3	348505	1	BX571870	Photorhab
44		177.6	41.8	6149	1	AY138456	Xenorhabd
44	c	169.6	39.9	308015	1	AB016783	Pseudomon
45		165.8	39.0	1450_8		CRU279023	Candida r

ALIGNMENTS

RESULT 1
BX640416
LOCUS   BX640416   349354 bp   DNA   linear   BCT 16-SEP-2003
DEFINITION Bordetella pertussis strain Tohama I, complete genome; segment
            6/12.
ACCESSION   BX640416 BX470248
VERSION   BX640416.1 GI:33572374
KEYWORDS   complete genome.
SOURCE   Bordetella pertussis Tohama I
  ORGANISM Bordetella pertussis Tohama I
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Alcaligenaceae; Bordetella.
REFERENCE   1
AUTHORS   Parkhill,J., Sebaihia,M., Preston,A., Murphy,L.D., Thomson,N.,
            Harris,D.E., Holden,M.T.G., Churcher,C.R., Bentley,S.D.,

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: September 20, 2004, 15:12:20 / Search time 1546.52 Seconds
                  (without alignments)
                  8206.421 Million cell updates/sec

Title: US-09-825-769A-11
Perfect score: 425
Sequence: 1 ctcgacctgcagaagctgaa..........ggtcggcctgcacgtcggc 425
Scoring table: IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0
Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :   BST:*
                  1: em_estba:*
                  2: em_esthum:*
                  3: em_estin:*
                  4: em_estmu:*
                  5: em_estov:*
                  6: em_estpl:*
                  7: em_estro:*
                  8: em_htc:*
                  9: em_est1:*
                  10: gb_est2:*
                  11: gb_htc:*
                  12: gb_est3:*
                  13: gb_est4:*
                  14: gb_est5:*
                  15: em_estfun:*
                  16: em_estom:*
                  17: em_gss_hum:*
                  18: em_gss_inv:*
                  19: em_gss_pln:*
                  20: em_gss_vrt:*
                  21: em_gss_fun:*
                  22: em_gss_man:*
                  23: em_gss_mus:*
                  24: em_gss_pro:*
                  25: em_gss_rod:*
                  26: em_gss_phg:*
                  27: em_gss_vrl:*
                  28: gb_gssI:*
                  29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query Match	length	DB	ID	Description
c	1	281.6	66.3	853	28	BZ576834	BZ576834 msh2_5124
c	2	280	65.9	1202	28	BZ568162	BZ568162 pacs2-164
	3	196.8	46.3	904	28	BZ554829	BZ554829 pacs1-60
	4	155.4	36.6	658	13	BU672399	BU672399 WHB3304_D
c	5	154.6	36.4	554	14	CF847417	CF847417 psHB043xP
	6	152.8	36.0	715	13	BX613550	BX613550 WS1_4_D08
	7	150.2	35.3	1148	28	BZ575496	BZ575496 msh2_4487
	8	142.2	33.5	806	14	CF402355	CF402355 RTWW1_19
c	9	141.6	33.3	622	10	BE195251	BE195251 HVSMEh008
	10	137.8	32.4	869	14	CB659834	CB659834 OSJNBd001B
	11	133.6	31.4	457	12	BG933522	BG933522 WS1_4_D08
c	12	130.4	30.7	1172	28	BZ585565	BZ585565 pa98401_2
c	13	130.2	30.6	980	11	CN508MEH	CN508MEH Single re
	14	128.4	30.2	773	13	BX870153	BX870153
	15	128.2	30.2	1491	11	AY107707	AY107707 Zea mays
	16	126.6	29.8	378	14	W06757	W06757 SMEST0033_S
	17	125	29.4	614	13	BQ295727	BQ295727 1091042A0
	18	124	29.2	2027	11	AK075575	AK075575 Mus muscu
	19	123.4	29.0	781	14	CF672902	CF672902 RTCNT1_74
	20	122.4	28.8	734	12	BG864405	BG864405 602796606
	21	122.4	28.8	738	12	BI146788	BI146788 602911867
	22	122.4	28.8	885	14	CA984300	CA984300 AGENCOURT
c	23	122.4	28.6	728	28	BH593957	BH593957 BOHPG01TR
	24	120.8	28.4	602	9	AL932800	AL932800 AL932800
	25	119.8	28.2	796	28	BZ435501	BZ435501 BONQH32TR
	26	119	28.0	367	28	CC143041	CC143041 NDL_12K22
	27	118.6	27.9	798	12	BG970636	BG970636 602835509
	28	118.2	27.8	809	12	BI655170	BI655170 603280972
	29	117.8	27.7	635	12	BZ521402	BZ521402 9-58 Stev
	30	117.2	27.6	607	12	BJ247332	BJ247332 BJ247332
	31	116.2	27.4	964	28	BZ573485	BZ573485 msh2_3220
	32	116.2	27.4	304	28	CC142377	CC142377 NDL_12K20
	33	115.2	27.1	304	10	AW496204	AW496204 up50b08.y
	34	114.6	27.0	1925	11	AK005724	AK005724 Mus muscu
	35	114.4	26.9	571	10	BE593476	BE593476 WS1_98_C1
	36	114.4	26.9	684	14	CF843243	CF843243 psHB024xG
	37	112.8	26.5	649	14	CD671950	CD671950 fg08h12.y
	38	112.8	26.5	901	13	BQ228630	BQ228630 AGENCOURT
	39	112.8	26.5	963	12	BG105831	BG105831 602312489
	40	112.8	26.5	963	14	CF619347	CF619347 AGENCOURT
	41	111.4	26.2	464	10	AW201019	AW201019 se96h10.y
	42	110	25.9	934	13	BU174000	BU174000 AGENCOURT
	43	109.6	25.8	689	13	BG290339	BG290339 602385962
	44	108.6	25.6	668	12	BG448280	BG448280 102402lB0
	45	107.4	25.3	468	29	CG045338	CG045338 PUID36TD

ALIGNMENTS

RESULT 1
BZ576834/c
LOCUS   BZ576834   853 bp   DNA   linear   GSS 17-DEC-2002
DEFINITION msh2_5124.y2 msh Pseudomonas aeruginosa genomic clone msh2_5124,
            genomic survey sequence.
ACCESSION   BZ576834
VERSION   BZ576834.1 GI:27211895
KEYWORDS   GSS.
SOURCE   Pseudomonas aeruginosa
  ORGANISM Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1 (bases 1 to 853)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE   Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
  COMMENT   Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES   Location/Qualifiers

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: September 20, 2004, 15:12:20 ; Search time 2488.36 Seconds
(without alignments)
9754.272 Million cell updates/sec

Title: US-09-825-769A-10
Perfect score: 560
Sequence: 1 ggcgcaagccgcgcgcgngn..........tgcagtgggccgccgcactga 560

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151516995 residues

Total number of hits satisfying chosen parameters: 0

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

6940544

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: gb_vrt:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

---

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	556	99.3	349354	1	BX640416	Bordetell
2	552.8	98.7	346274	1	BX640443	Bordetell
3	548	97.9	344321	1	BX640429	Bordetell
4	396	70.7	193050	1	AL646062	Ralstonia
5	386.4	69.0	301846	1	AE016913	Chromobac
6	325.8	58.1	14652	1	AE004799	Pseudomon
7	308	55.0	7886	1	AF010139	Azotobact
8	300	53.6	1215	6	AX018964	Sequence
9	300	53.6	1215	6	BD136659	Process f
10	300	53.6	1233	6	BD073376	Process f
11	300	53.6	1251	6	AR387886	Sequence
12	300	53.6	3481	6	AX000480	Sequence
13	300	53.6	3481	6	BD073380	Process f
14	300	53.6	3810	6	AX000476	Sequence
15	300	53.6	3810	6	BD073378	Process f
16	300	53.6	9356	1	AE005483	Escherich
17	300	53.6	10614	1	AE000339	Escherich
18	300	53.6	10614	6	AX370241	Sequence
19	300	53.6	13332	2	D90883	E.coli geno
20	300	53.6	241679	2	AC020861	Mus muscu
21	300	53.6	270365	6	AC002562	Escherich
22	299.6	53.5	1232	6	AX000472	Sequence
23	298.4	53.3	300099	1	AE016764	Escherich
24	296.8	53.0	9329	1	AE015273	Shigella
25	296.8	53.0	290380	1	AE016987	Shigella
26	290.4	51.9	10203	1	AE002486	Neisseria
27	290.4	51.9	311249	1	AE016860	Pseudomon
28	290.4	51.9	349980	6	AX044032	Sequence
29	288.8	51.6	145050	1	AL627275	Salmonell
30	288.8	51.6	300431	1	AE016835	Salmonell
31	288.8	51.6	329861	1	NMA522491	Neisseria
32	287.2	51.3	19982	1	AL162756	Neisseria
33	279.2	49.9	300861	1	AE008815	Salmonell
34	252	45.0	14261	1	AE016777	Pseudomon
35	248.8	44.4	11059	1	AE004160	Vibrio ch
36	248.8	44.4	13584	1	AE015668	Shewanell
37	248.8	44.4	204050	1	AE013736	Yersinia
38	247.2	44.1	251720	1	AJ414154	Yersinia
39	240.8	43.0	301235	1	AP005075	Vibrio vu
40	240.8	43.0	301442	1	AE016798	Vibrio vu
41	240.8	43.0	348505	1	BX571870	Photorhab
42	220	39.3	6149	1	AY138456	Xenorhabd
43	220	39.3	10507	1	U32721	Haemophilus
44	220	39.3	11223	1	AE006068	Pasteurel
45	220	39.3	110000	6	AR274513_03	Continuation (4 of

ALIGNMENTS

RESULT 1
BX640416
LOCUS   BX640416   349354 bp   DNA   linear   BCT 16-SEP-2003
DEFINITION Bordetella pertussis strain Tohama I, complete genome; segment
            6/12.
ACCESSION   BX640416 BX470248
VERSION   BX640416.1 GI:33572374
KEYWORDS   complete genome.
SOURCE   Bordetella pertussis Tohama I
ORGANISM   Bordetella pertussis Tohama I
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Alcaligenaceae; Bordetella.
REFERENCE   1
AUTHORS   Parkhill,J., Sebaihia,M., Preston,A., Murphy,L.D., Thomson,N.,
            Harris,D.B., Holden,M.T.G., Churcher,C.R., Bentley,S.D.,

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: September 20, 2004, 15:12:19 ; Search time 302.234 Seconds
(without alignments)
7871.356 Million cell updates/sec

Title: US-09-825-769A-10
Perfect score: 560
Sequence: 1 ggcgcaagccgcgcgtgngn................tgcagtgggccgcgcactga 560

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	ID	Description
1	556	99.3	560	6	AAS15488	Aas15488 B. pertus
2	556	99.3	1212	6	AAS15484	Aas15484 DNA encod
3	353.4	63.1	1212	6	ABG90070	Abg90070 M. capsul
4	334	59.6	425	6	AAS15487	Aas15487 B. pertus
5	308.8	55.1	1212	7	ABZ39842	Abz39842 N. gonorr
6	300	53.6	1215	2	AAZ09786	Aaz09786 B. coli b
7	300	53.6	3481	2	AAX02815	Aax02815 DNA encod
8	300	53.6	3810	2	AAX02813	Aax02813 DE1973127
9	300	53.6	10614	4	AAS46259	Aas46259 DNA encod
c 10	299.6	53.5	1232	3	AAX02811	Aax02811 B. coli b
c 11	290.4	51.9	13786	3	AAA81480	Aaa81480 N. mening
c 12	290.4	51.9	110000	3	AAB81490_13	Aab81490_13
c 13	290.4	51.9	349980	3	AAF21610	Aaf21610 Neisseria
c 14	254.2	45.4	3414	5	AAS73033	Aas73033 DNA encod
15	240.8	43.0	1215	7	ACF72052	Acf72052 Photorhab
16	240.8	43.0	110000	7	ACF67367_52	Acf67367_52
17	240.8	43.0	110000	7	ACF65387_4	Acf65387_4
c 18	220	39.3	110000	2	ACF42063_03	Acf42063_03
19	213.6	38.1	6209	4	AAF28545	Aaf28545 Genomic f
20	205	36.6	1548	5	AAS88547	Aas88547 DNA encod
21	205	36.6	1548	5	AAS93345	Aas93345 DNA encod
22	205	36.6	1548	5	AAS73075	Aas73075 DNA encod
23	202.4	36.1	1260	8	ADA32431	Ada32431 DNA encod

RESULT 1
AAS15488

ID AAS15488 standard; DNA; 560 BP.
XX
AC AAS15488;
XX
DT 29-AUG-2003 (revised)
DT 14-FEB-2002 (first entry)
XX
DE B. pertussis cysteine desulphinase DNA (DSFS3612) from strain BP536.
XX
KW Large-scale bacterial toxin production; pertussis toxin; PT;
KW whooping cough vaccine; cysteine desulphinase; ds.
XX
OS Bordetella pertussis; strain BP536.
XX
PN WO200174862-A2.
XX
PD 11-OCT-2001.
XX
PF 04-APR-2001; 2001WO-US010938.
XX
PR 04-APR-2000; 2000US-0194478P.
PR 04-APR-2000; 2000US-0194482P.
XX
PA (BAXT) BAXTER INT INC.
PA (BAXT-) BAXTER HEALTHCARE SA.
XX
PI Blake MS, Bogdan JA, Nazario-Larrieu J;
XX
PS Disclosure; Fig 7; 46pp; English.
XX
CC The present invention relates to a method of enhancing the production of
CC bacterial toxins in large-scale cultures. The method comprises
CC eliminating or reducing toxin expression inhibitors formed by toxin
CC producing bacteria. The invention provides a method for producing
CC pertussis toxin (PT) from Bordetella pertussis by introducing a soluble
CC salt into the growth medium that sequesters sulphate and/or employing a
B. pertussis cysteine desulphinase knockout mutant. The method is useful

WPI; 2002-010777/01.

Enhancing production of bacterial toxins comprises eliminating or
reducing toxin expression inhibitors formed by toxin producing bacteria
by adding at least one soluble metal salt that forms an insoluble complex
with sulfate.

DR WO20017486A-A2.

ALIGNMENTS

24	192.6	34.4	710	7	ACF66682	Acf66682 Photorhab
25	187.6	33.5	1554	7	ABT20720	Abt20720 Aspergill
26	184.6	33.0	1518	4	ABL09101	Abl09101 Drosophil
27	184.6	33.0	3041	4	ABL11656	Abl11656 Drosophil
28	184.6	33.0	3576	4	ABL09100	Abl09100 Drosophil
29	173	30.9	1649	3	AAF18102	Aaf18102 Lung canc
30	173	30.9	2380	4	AAH14339	Aah14339 Human cDN
31	173	30.9	3224	5	ABV27901	Abv27901 Human pro
32	170.8	30.5	1434	6	ADB69794	Adb69794 C. neofor
33	162.6	29.0	2092	4	AAH14100	Aah14100 Arabidops
34	149.4	26.7	1362	3	ABZ12797	Abz12797 A. thalia
35	149.4	26.7	1399	3	AAZ98363	Aaz98363 Arabidops
36	149.4	26.7	1656	3	AAC42345	Aac42345 Human ova
37	142.6	25.5	495		ABI80979	Abi80979 Human cDN
38	140.2	25.0	40681	6	ABA92787_6	Continuation (7 of
39	137	24.5	1778	7	ABT20122	Abt20122 Aspergill
40	137	24.5	3140	7	ABT17712	Abt17712 Aspergill
41	137	24.5	3778	7	ABT19526	Abt19526 Aspergill
42	127.4	22.8	1494	7	AAD48348	Aad48348 Saccharom
43	127.4	22.8	11198	4	AAD07496	Aad07496 pYIR110P
44	127.4	22.8	11201	4	AAD07498	Aad07498 pYMR107P
45	127.4	22.8	11204	4	AAD07499	Aad07499 pZBO1P DN

OM nucleic - nucleic search, using sw model

Run on: September 20, 2004, 15:12:20 ; Search time 2037.77 Seconds

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

(without alignments)
8206.421 Million cell updates/sec

Title: US-09-825-769A-10
Perfect score: 560
Sequence: 1 ggcgcaagccgcgccgtgngn..........tgcagtggccgcgcactga 560

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2751289 seqs, 1493109090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
BST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_estfun:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_rod:*
26: em_gss_pro:*
27: em_gss_phg:*
28: em_gss_vrl:*
29: gb_gss1:*
30: gb_gss2:*

c	Result No.	Score	Query Match Length	DB	ID	Description
c	1	291	52.0	853	28	BZ576834	BZ576834 msh2_5124
c	2	290.2	51.8	1202	28	BZ568162	BZ568162 pacs2-164
	3	223.2	39.9	367	28	CC143041	CC143041 NDL.12K22
	4	220	39.3	684	14	CF843243	CF843243 psHB024xG

c	No.	Score	%	Length	DB	ID	ID / Description
	5	202.2	36.1	869	14	CB659834	CB659834 OSJNBd01B
	6	194.4	34.7	304	28	CC142377	CC142377 NDL.12K20
	7	181	32.3	681	13	BX253196	BX253196 BX253196
	8	181	32.3	806	14	CF402355	CF402355 RTWW1_19_
	9	179.4	32.0	1167	29	BI146788	BI146788 602911867
	10	179.4	32.0	2027	11	AK075575	AK075575 Mus muscu
	11	178.2	31.8	1491	11	AY107707	AY107707 Zea mays
	12	177.4	31.7	612	9	AI455290	AI455290 LD226661.3
	13	176.2	31.5	734	12	BG864405	BG864405 602798606
	14	175.2	31.3	885	14	CA984300	CA984300 AGENCOURT
	15	174	31.1	738	12	BI146788	BI146788 602911867
	16	173.6	31.0	803	14	CA267465	CA267465 SCJLLB207
	17	173	30.9	801	12	BG743654	BG743654 602632823
	18	173	30.9	885	14	CA454242	CA454242 AGENCOURT
	19	171.4	30.6	1167	29	AY417651	AY417651 Homo sapi
	20	170.4	30.4	947	13	CF409485	CF409485 CH3#060_H
	21	170.4	30.4	885	14	BU749145	BU749145 CH3#025_A
	22	170	30.4	408	14	CF841038	CF841038 psHB012xI
	23	170	30.4	1925	11	AK005724	AK005724 Mus muscu
	24	169.4	30.3	594	14	CF354228	CF354228 1ac20d09.
	25	169.4	30.3	521	13	CA163913	CA163913 SCRUR2308
	26	169	30.2	663	14	CD920753	CD920753 G608.1118B
	27	164.4	29.4	665	10	AW496203	AW496203 up50b08.Y
	28	163.2	29.1	669	9	AJ432663	AJ432663 OSJNBa06I
	29	163.2	29.1	760	14	CB672483	CB672483 OSJNBa06I
c	30	161.6	28.9	645	13	BQ766418	BQ766418 EBroo8_SQ
c	31	161.6	28.9	801	14	CB659833	CB659833 OSJNEd01B
c	32	160.8	28.7	723	9	AV911459	AV911459 AV911459
c	33	160.6	28.7	658	13	BU672399	BU672399 WHB3304_D
c	34	157.6	28.1	567	12	BJ297418	BJ297418 BJ297418
c	35	157	28.0	555	13	CA009764	CA009764 HT12P02u
c	36	156.4	27.9	512	14	CD922112	CD922112 G750.101C
c	37	154.2	27.5	809	12	BI655170	BI655170 603280972
c	38	153.6	27.4	517	9	AV437846	AV437846 AV437846
c	39	152.6	27.2	980	11	BX018453	BX018453 Single re
c	40	149	26.6	868	14	CK266135	CK266135 BST712213
c	41	148.6	26.5	744	14	CF672830	CF672830 RTCNT1_74
c	42	147.4	26.3	520	13	BX253508	BX253508 BX253508
c	43	146	26.1	614	13	BQ295727	BQ295727 1091042A0
c	44	144.2	25.8	677	13	BU060836	BU060836 Pgr-C_1_K
c	45	144.2	25.8	942	14	CK245065	CK245065 BST728702

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
BZ576834/c
LOCUS   BZ576834   853 bp   DNA   linear   GSS 17-DEC-2002
DEFINITION msh2_5124.y2 msh Pseudomonas aeruginosa genomic clone msh2_5124,
            genomic survey sequence.
ACCESSION   BZ576834
VERSION   BZ576834.1 GI:27211895
KEYWORDS   GSS.
SOURCE   Pseudomonas aeruginosa
  ORGANISM Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1 (bases 1 to 853)
  AUTHORS   Spencer,D.H., Kaul,R. and Olsen,M.V.
            Burns,J.L., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
  TITLE   Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT   Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES   Location/Qualifiers

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: September 20, 2004, 15:12:20 ; Search time 2055.97 seconds
(without alignments)
8206.421 Million cell updates/sec

Title: US-09-825-769A-9
Perfect score: 565
Sequence: 1 atgagcaatcgcccatcta..........ccaacggcaagtcgagatc 565

Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0
Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Databases :
  BST:*
  1: em_estba:*
  2: em_esthum:*
  3: em_estln:*
  4: em_estmu:*
  5: em_estov:*
  6: em_estpl:*
  7: em_estro:*
  8: em_htc:*
  9: em_est1:*
  10: gb_est2:*
  11: gb_hc:*
  12: gb_est3:*
  13: gb_est4:*
  14: gb_est5:*
  15: em_estfun:*
  16: em_estom:*
  17: em_est_hum:*
  18: em_gss_inv:*
  19: em_gss_pln:*
  20: em_gss_vrt:*
  21: em_gss_fun:*
  22: em_gss_mam:*
  23: em_gss_mus:*
  24: em_gss_pro:*
  25: em_gss_rod:*
  26: em_gss_phg:*
  27: em_gss_vrl:*
  28: gb_gss1:*
  29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	284	50.3	876	28	BZ577360	BZ577360 msh2_5370
2	265.8	47.0	863	28	BZ562611	BZ562611 pacs2-164
3	245	43.4	1071	28	BZ564477	BZ564477 pacs2-164
4	233.2	41.3	674	13	CA071128	CA071128 SCCCAM100
5	233.2	41.3	692	13	CA071140	CA071140 SCCCAM100
6	225.2	39.9	795	14	CB672482	CB672482 OSSNBe061
7	221.4	39.2	655	14	CA186265	CA186265 SCSGST311
8	208.2	37.9	695	9	AU161333	AU161333 AU161333
9	214.2	37.9	862	13	BX466420	BX466420
10	212	37.5	668	13	BQ840710	BQ840710 WHB4202_P
11	211.6	37.5	474	13	CA012321	CA012321 HT05B16T
12	210.2	37.2	673	12	BI997818	BI997818 1031052A1
13	208.8	37.0	651	13	BX624227	BX624227
14	208.2	36.8	941	11	AU184452	AU184452 Single_re
15	206.4	36.5	759	13	BX627075	BX627075 BX627075
16	205	36.3	742	10	BF496153	BF496153
17	205	36.3	746	9	AI516994	AI516994 GH27486.5
18	201.2	35.6	628	11	AY107707	AY107707 Zea mays
19	200.2	35.5	434	13	BX767041	BX767041
20	199	35.2	690	13	BX622505	BX622505
21	197.4	34.9	655	14	CA214446	CA214446 SCUTFL105
22	196.4	34.8	612	9	CD036744	CD036744 mgsu011xD
23	196.4	34.8	612	9	AU091715	AU091715 AU091715
24	194.4	34.4	653	12	BI142155	BI142155 SD16161.5
25	193	34.2	654	13	BU122793	BU122793 603148834
26	192.2	34.0	864	14	CP377821	CP377821 AGENCOURT
27	187.6	33.2	434	9	AU093662	AU093662 AGENCOURT
28	184.6	32.7	739	14	CF285833	CF285833 AGENCOURT
29	184.4	32.6	682	9	AJ456871	AJ456871 AJ456871
30	184.2	32.6	864	12	BU612099	BU612099 BU612099
31	184.2	32.6	917	13	BU175049	BU175049 AGENCOURT
32	183.4	32.5	725	14	CF450610	CF450610 EST686955
33	183.2	32.4	633	14	CB577181	CB577181 AMGNNDC:N
34	183.2	32.4	744	12	BI330589	BI330589 602981030
35	178.6	31.6	750	12	BI854562	BI854562 603380275
36	178.6	31.6	803	10	BF608290	BF608290 MY1_00123
37	178.6	31.6	829	14	CF837753	CF837753 UCRCS03_0
38	178.6	31.6	854	14	CF586180	CF586180 AGENCOURT
39	178.6	31.6	864	12	BI831420	BI831420 603074521
40	178.6	31.6	774	12	BI763016	BI763016 603047806
41	178.6	31.6	876	12	BI824722	BI824722 603033891
42	178.4	31.6	1018	12	BM452489	BM452489 AGENCOURT
43	178.4	31.6	1028	13	BQ073152	BQ073152 AGENCOURT
44	178.4	31.6	1080	12	BM808412	BM808412 AGENCOURT
45	178.4	31.6	1134	12	BM545535	BM545535 AGENCOURT

ALIGNMENTS

RESULT 1
BZ577360/c
LOCUS   BZ577360   876 bp   DNA   linear   GSS 17-DEC-2002
DEFINITION msh2_5370.y2 msh Pseudomonas aeruginosa genomic clone msh2_5370,
            genomic survey sequence.
ACCESSION   BZ577360
VERSION   BZ577360.1 GI:27212421
KEYWORDS   GSS.
SOURCE   Pseudomonas aeruginosa
ORGANISM   Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1 (bases 1 to 876)
AUTHORS   Spencer,D.H., Raymond,C.K., Smith,B.B., Sims,B.B., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE   Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
JOURNAL   J. Bacteriol. (2002) In press
COMMENT   Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES   Location/Qualifiers

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on)          September 20, 2004, 15:12:19 ; Search time 304,932 seconds
                                       (without alignments)
                                       7871.356 Million cell updates/sec

Title:           US-09-825-769A-9
Perfect score:   565
Sequence:        1 atgagcaatcgcccatcta.........ccaacggcaaggtcgagatc 565

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters:        6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_29Jan04:*
                 1: geneseqn1980s:*
                 2: geneseqn1990s:*
                 3: geneseqn2000s:*
                 4: geneseqn2001as:*
                 5: geneseqn2001bs:*
                 6: geneseqn2002s:*
                 7: geneseqn2003as:*
                 8: geneseqn2003bs:*
                 9: geneseqn2003cs:*
                10: geneseqn2004s:*   .

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

```
Result             %
No.    Score  Query Match Length DB ID       Description
-----  -----  ---------------------------    -----------
  1    565    100.0  565    6  AAS15485      Aas15485 B. pertus
  2    539.8   95.5 1212    6  AAS15484      Aas15484 DNA encod
  3    305     54.0 1212    6  ABQ90070      Abq90070 M. capsul
  4    275.2   48.7 1212    7  ABZ39842      Abz39842 N. gonorr
c 5    260     46.0 13786   3  AAA81480      Aaa81480 N. mening
  6    260     46.0 110000  3  AAA81490_13   Continuation (14 o
  7    260     46.0 349980  3  AAF21610      Aaf21610 Neisseria
  8    250     44.2 1215    2  AAZ09786      Aaz09786 E. coli b
  9    250     44.2 1232    2  AAX02811      Aax02811 E. coli b
c 10   250     44.2 3414    5  AAS73033      Aas73033 DNA encod
c 11   250     44.2 3481    2  AAX02815      Aax02815 DEl973127
c 12   250     44.2 3810    2  AAX02813      Aax02813 DEl973127
c 13   250     44.2 10614   4  AAS46259      Aas46259 DNA encod
  14   250     44.2 1215    7  ACF72052      Acf72052 Photorhab
  15   209     37.0 110000  7  ACF65387_52   Continuation (53 o
  16   209     37.0 110000  7  ACF65387_4    Continuation (5 of
  17   205     36.3 1518    4  ABL09101      Abl09101 Drosophil
  18   205     36.3 3576    4  ABL09100      Abl09100 Drosophil
  19   197     34.9 1434    9  ADB69794      Adb69794 C. neofor
  20   197     34.9 1653    9  ADB69433      Adb69433 C. neofor
  21   197     34.9 3653    9  ADB69072      Adb69072 C. neofor
c 22   182.6   32.3 1419    5  AAS88509      Aas88509 DNA encod
  23   178.4   31.6 2092    4  AAH14100      Aah14100 Human CDN
  24   178.4   31.6 2380    4  AAH14339      Aah14339 Human CDN
  25   178.4   31.6 3224    5  ABV27901      Abv27901 Human pro
  26   176.8   31.3 1100000 2  AAT42063_03   Continuation (4 of
  27   170.4   30.2  863    7  AAH05422      Aah05422 Human cDN
  28   170.2   30.1 1140    7  ABT18306      Abt18306 Aspergill
  29   170.2   30.1 1140    7  ABT18900      Abt18900 Aspergill
  30   170.2   30.1 1554    7  ABT20720      Abt20720 Aspergill
  31   170.2   30.1 1778    7  ABT20122      Abt20122 Aspergill
  32   170.2   30.1 3140    7  ABT17712      Abt17712 Aspergill
  33   170.2   30.1 3778    7  ABT19526      Abt19526 Aspergill
  34   169.8   30.1 1260    8  ADA32431      Ada32431 DNA encod
  35   169.8   30.0 62909   4  AAF28545      Aaf28545 Genomic f
  36   156.6   27.7 1362    6  ABZ12797      Abz12797 Arabidops
  37   156.6   27.7 1399    4  AAZ98363      Aaz98363 A. thalia
  38   156.6   27.7 1656    3  AAC42345      Aac42345 Arabidops
  39   155.8   27.6 1494    7  AAD48348      Aad48348 Saccharom
  40   155.8   27.6 11198   4  AAD07496      Aad07496 pYLR110P
  41   155.8   27.6 11201   4  AAD07498      Aad07498 pYMR107P
  42   155.8   27.6 11204   4  AAD07499      Aad07499 p2EO1P DN
  43   155.8   27.6 11427   4  AAD07497      Aad07497 pYMR251AP
  44   155.8   27.6 12008   4  AAD07500      Aad07500 pPRIBP DN
  45   155.8   27.6 12844   4  AAD07492      Aad07492 pYLR110P+
```

ALIGNMENTS

```
RESULT 1
AAS15485
ID   AAS15485 standard; DNA; 565 BP.
XX
AC   AAS15485;
XX
DT   29-AUG-2003  (revised)
DT   14-FEB-2002  (first entry)
XX
DE   B. pertussis cysteine desulphinase DNA (DSF536F1) from strain BP536.
XX
KW   Large-scale bacterial toxin production; pertussis toxin; PT;
KW   whooping cough vaccine; cysteine desulphinase; ds.
XX
OS   Bordetella pertussis; strain BP536.
XX
PN   WO200174862-A2.
XX
PD   11-OCT-2001.
XX
PF   04-APR-2001; 2001WO-US010938.
XX
PR   04-APR-2000; 2000US-0194478P.
PR   04-APR-2000; 2000US-0194482P.
XX
PA   (BAXT ) BAXTER INT INC.
PA   (BAXT-) BAXTER HEALTHCARE SA.
XX
PI   Blake MS,  Bogdan JA,  Nazario-Larrieu J;
XX
DR   WPI; 2002-010777/01.
XX
PT   Enhancing production of bacterial toxins comprises eliminating or
PT   reducing toxin expression inhibitors formed by toxin producing bacteria
PT   by adding at least one soluble metal salt that forms an insoluble complex
PT   with sulfate.
XX
PS   Disclosure; Fig 7; 46pp; English.
XX
CC   The present invention relates to a method of enhancing the production of
CC   bacterial toxins in large-scale cultures. The method comprises
CC   eliminating or reducing toxin expression inhibitors formed by toxin
CC   producing bacteria. The invention provides a method for producing
CC   pertussis toxin (PT) from Bordetella pertussis by introducing a soluble
CC   salt into the growth medium that sequesters sulphate and/or employing a
CC   B. pertussis cysteine desulphinase knockout mutant. The method is useful
```